OIPE

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/09/939,537   TIME: 09:33:06

Input Set : N:\Crf3\RULE60\09939537.txt
Output Set: N:\CRF3\02112002\I939537.raw

SEQUENCE LISTING
```
C--> 5 (1) GENERAL INFORMATION:
     7      (i) APPLICANT: Seed, Brian
     8                     Banapour, Babak
     9                     Romeo, Charles
    10                     Kolanus, Waldemar
C--> 12    (ii) TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
    13                     CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
    15   (iii) NUMBER OF SEQUENCES: 53
    17    (iv) CORRESPONDENCE ADDRESS:
    18         (A) ADDRESSEE: Clark & Elbing LLP
    19         (B) STREET: 176 Federal Street
    20         (C) CITY: Boston
    21         (D) STATE: MA
    22         (E) COUNTRY: USA
    23         (F) ZIP: 02110
    25     (v) COMPUTER READABLE FORM:
    26         (A) MEDIUM TYPE: Diskette
    27         (B) COMPUTER: IBM Compatible
    28         (C) OPERATING SYSTEM: DOS
    29         (D) SOFTWARE: FastSEQ for Windows Version 2.0
    31    (vi) CURRENT APPLICATION DATA:
C--> 32         (A) APPLICATION NUMBER: US/09/939,537
C--> 33         (B) FILING DATE: 24-Aug-2001
    39         (C) CLASSIFICATION:
C--> 41   (vii) PRIOR APPLICATION DATA:
    37         (A) APPLICATION NUMBER: 08/284,391
    38         (B) FILING DATE: 02-AUG-1994
    42         (A) APPLICATION NUMBER: 08/195,395
    43         (B) FILING DATE: 14-FEB-1994
    45         (A) APPLICATION NUMBER: 07/847,566
    46         (B) FILING DATE: 06-MAR-1992
    48         (A) APPLICATION NUMBER: 07/665,961
    49         (B) FILING DATE: 07-MAR-1991
    51  (viii) ATTORNEY/AGENT INFORMATION:
    52         (A) NAME: Elbing, Karen L
    53         (B) REGISTRATION NUMBER: 35,238
    54         (C) REFERENCE/DOCKET NUMBER: 00786/247001
    56    (ix) TELECOMMUNICATION INFORMATION:
    57         (A) TELEPHONE: 617-428-0200
    58         (B) TELEFAX: 617-428-7045
    59         (C) TELEX:
    62 (2) INFORMATION FOR SEQ ID NO: 1:
```

ENTERED

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/09/939,537   TIME: 09:33:06

     Input Set : N:\Crf3\RULE60\09939537.txt
     Output Set: N:\CRF3\02112002\I939537.raw

```
 64      (i) SEQUENCE CHARACTERISTICS:
 65          (A) LENGTH: 1728 base pairs
 66          (B) TYPE: nucleic acid
 67          (C) STRANDEDNESS: double
 68          (D) TOPOLOGY: linear
 70      (ii) MOLECULE TYPE: cDNA
 72      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 74  ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC GCTCCTCCCA     60
 75  GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG GGGATACAGT GGAACTGACC    120
 76  TGTACAGCTT CCCAGAAGAA GAGCATACAA TTCCACTGGA AAAACTCCAA CCAGATAAAG    180
 77  ATTCTGGGAA ATCAGGGCTC CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT    240
 78  GACTCAAGAA GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG    300
 79  ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA GGTGCAATTG    360
 80  CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC TTCAGGGGCA GAGCCTGACC    420
 81  CTGACCTTGG AGAGCCCCCC TGGTAGTAGC CCCTCAGTGC AATGTAGGAG TCCAAGGGGT    480
 82  AAAAACATAC AGGGGGGGAA GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC    540
 83  ACCTGGACAT GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG    600
 84  GTGCTAGCTT CCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA ACAGGTGGAG    660
 85  TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG GCAGTGGCGA GCTGTGGTGG    720
 86  CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT TGGATCACCT TTGACCTGAA GAACAAGGAA    780
 87  GTGTCTGTAA AACGGGTTAC CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC    840
 88  CACCTCACCC TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC    900
 89  CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT GAGAGCCACT    960
 90  CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA CCTCCCCTAA GCTGATGCTG   1020
 91  AGCTTGAAAC TGGAGAACAA GGAGGCAAAG GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG   1080
 92  CTGAACCCTG AGGCGGGGAT GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCCTGCTG   1140
 93  GAATCCAACA TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCCAAACTC   1200
 94  TGCTACTTGC TAGATGGAAT CCTCTTCATC TACGGAGTCA TCATCACAGC CCTGTACCTG   1260
 95  AGAGCAAAAT TCAGCAGGAG TGCAGAGACT GCTGCCAACC TGCAGGACCC CAACCAGCTC   1320
 96  TACAATGAGC TCAATCTAGG GCGAAGAGAG GAATATGACG TCTTGGAGAA GAAGCGGGCT   1380
 97  CGGGATCCAG AGATGGGAGG CAAACAGCAG AGGAGGAGGA ACCCCCAGGA AGGCGTATAC   1440
 98  AATGCACTGC AGAAAGACAA GATGCCAGAA GCCTACAGTG AGATCGGCAC AAAAGGCGAG   1500
 99  AGGCGGAGAG GCAAGGGGCA CGATGGCCTT TACCAGGACA GCCACTTCCA AGCAGTGCAG   1560
100  TTCGGGAACA GAAGAGAGAG AGAAGGTTCA GAACTCACAA GGACCCTTGG GTTAAGAGCC   1620
101  CGCCCCAAAG GTGAAAGCAC CCAGCAGAGT AGCCAATCCT GTGCCAGCGT CTTCAGCATC   1680
102  CCCACTCTGT GGAGTCCATG GCCACCCAGT AGCAGCTCCC AGCTCTAA              1728
104 (2) INFORMATION FOR SEQ ID NO: 2:
106      (i) SEQUENCE CHARACTERISTICS:
107          (A) LENGTH: 1389 base pairs
108          (B) TYPE: nucleic acid
109          (C) STRANDEDNESS: double
110          (D) TOPOLOGY: linear
112      (ii) MOLECULE TYPE: cDNA
114      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
116  ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC GCTCCTCCCA     60
117  GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG GGGATACAGT GGAACTGACC    120
118  TGTACAGCTT CCCAGAAGAA GAGCATACAA TTCCACTGGA AAAACTCCAA CCAGATAAAG    180
119  ATTCTGGGAA ATCAGGGCTC CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT    240
120  GACTCAAGAA GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG    300
```

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/09/939,537   TIME: 09:33:06

Input Set : N:\Crf3\RULE60\09939537.txt
Output Set: N:\CRF3\02112002\I939537.raw

```
121  ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA GGTGCAATTG   360
122  CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC TTCAGGGGCA GAGCCTGACC   420
123  CTGACCTTGG AGAGCCCCCC TGGTAGTAGC CCCTCAGTGC AATGTAGGAG TCCAAGGGGT   480
124  AAAAACATAC AGGGGGGGAA GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC   540
125  ACCTGGACAT GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG   600
126  GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA ACAGGTGGAG   660
127  TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG GCAGTGGCGA GCTGTGGTGG   720
128  CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT TGGATCACCT TTGACCTGAA GAACAAGGAA   780
129  GTGTCTGTAA AACGGGTTAC CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC   840
130  CACCTCACCC TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC   900
131  CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT GAGAGCCACT   960
132  CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA CCTCCCCTAA GCTGATGCTG  1020
133  AGCTTGAAAC TGGAGAACAA GGAGGCAAAG GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG  1080
134  CTGAACCCTG AGGCGGGGAT GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCCTGCTG  1140
135  GAATCCAACA TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCGCAGCTC  1200
136  TGCTATATCC TGGATGCCAT CCTGTTTTTG TATGGTATTG TCCTTACCCT GCTCTACTGT  1260
137  CGACTCAAGA TCCAGGTCCG AAAGGCAGAC ATAGCCAGCC GTGAGAAATC AGATGCTGTC  1320
138  TACACGGGCC TGAACACCCG GAACCAGGAG ACATATGAGA CTCTGAAACA TGAGAAACCA  1380
139  CCCCAATAG                                                         1389
141 (2) INFORMATION FOR SEQ ID NO: 3:
143       (i) SEQUENCE CHARACTERISTICS:
144            (A) LENGTH: 1599 base pairs
145            (B) TYPE: nucleic acid
146            (C) STRANDEDNESS: double
147            (D) TOPOLOGY: linear
149       (ii) MOLECULE TYPE: cDNA
151       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
153  ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC GCTCCTCCCA    60
154  GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG GGGATACAGT GGAACTGACC   120
155  TGTACAGCTT CCCAGAAGAA GAGCATACAA TTCCACTGGA AAAACTCCAA CCAGATAAAG   180
156  ATTCTGGGAA ATCAGGGCTC CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT   240
157  GACTCAAGAA GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG   300
158  ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA GGTGCAATTG   360
159  CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC TTCAGGGGCA GAGCCTGACC   420
160  CTGACCTTGG AGAGCCCCCC TGGTAGTAGC CCCTCAGTGC AATGTAGGAG TCCAAGGGGT   480
161  AAAAACATAC AGGGGGGGAA GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC   540
162  ACCTGGACAT GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG   600
163  GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA ACAGGTGGAG   660
164  TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG GCAGTGGCGA GCTGTGGTGG   720
165  CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT TGGATCACCT TTGACCTGAA GAACAAGGAA   780
166  GTGTCTGTAA AACGGGTTAC CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC   840
167  CACCTCACCC TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC   900
168  CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT GAGAGCCACT   960
169  CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA CCTCCCCTAA GCTGATGCTG  1020
170  AGCTTGAAAC TGGAGAACAA GGAGGCAAAG GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG  1080
171  CTGAACCCTG AGGCGGGGAT GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCCTGCTG  1140
172  GAATCCAACA TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCCAAACTC  1200
173  TGCTACCTGC TGGATGGAAT CCTCTTCATC TATGGTGTCA TTCTCACTGC CTTGTTCCTG  1260
174  AGAGTGAAGT TCAGCAGGAG CGCAGAGCCC CCCGCGTACC AGCAGGGCCA GAACCAGCTC  1320
```

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/09/939,537   TIME: 09:33:06

Input Set : N:\Crf3\RULE60\09939537.txt
Output Set: N:\CRF3\02112002\I939537.raw

```
175  TATAACGAGC TCAATCTAGG ACGAAGAGAG GAGTACGATG TTTTGGACAA GAGACGTGGC  1380
176  CGGGACCCTG AGATGGGGGG AAAGCCGAGA AGGAAGAACC CTCAGGAAGG CCTGTACAAT  1440
177  GAACTGCAGA AAGATAAGAT GGCGGAGGCC TACAGTGAGA TTGGGATGAA AGGCGAGCGC  1500
178  CGGAGGGGCA AGGGGCACGA TGGCCTTTAC CAGGGTCTCA GTACAGCCAC CAAGGACACC  1560
179  TACGACGCCC TTCACATGCA GGCCCTGCCC CCTCGCTAA                         1599
181  (2) INFORMATION FOR SEQ ID NO: 4:
183       (i) SEQUENCE CHARACTERISTICS:
184           (A) LENGTH: 575 amino acids
185           (B) TYPE: amino acid
186           (C) STRANDEDNESS: single
187           (D) TOPOLOGY: linear
189      (ii) MOLECULE TYPE: protein
191      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
193  Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu
194   1               5                  10                  15
195  Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys
196              20                  25                  30
197  Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser
198          35                  40                  45
199  Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn
200      50                  55                  60
201  Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala
202  65                  70                  75                  80
203  Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile
204              85                  90                  95
205  Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu
206              100                 105                 110
207  Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn
208              115                 120                 125
209  Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu
210          130                 135                 140
211  Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly
212  145                 150                 155                 160
213  Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu
214              165                 170                 175
215  Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys
216              180                 185                 190
217  Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
218              195                 200                 205
219  Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro
220      210                 215                 220
221  Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
222  225                 230                 235                 240
223  Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu
224              245                 250                 255
225  Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu
226              260                 265                 270
227  Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu
228          275                 280                 285
```

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/09/939,537   TIME: 09:33:06

Input Set : N:\Crf3\RULE60\09939537.txt
Output Set: N:\CRF3\02112002\I939537.raw

```
229  Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys
230      290             295             300
231  Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Ala Thr
232  305             310             315             320
233  Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro
234              325             330             335
235  Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser
236      ·340            345             350
237  Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Met Trp
238      355             360             365
239  Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile
240      370             375             380
241  Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Lys Leu
242  385             390             395             400
243  Cys Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Ile Thr
244              405             410             415
245  Ala Leu Tyr Leu Arg Ala Lys Phe Ser Arg Ser Ala Glu Thr Ala Ala
246              420             425             430
247  Asn Leu Gln Asp Pro Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg
248      435             440             445
249  Arg Glu Glu Tyr Asp Val Leu Glu Lys Lys Arg Ala Arg Asp Pro Glu
250      450             455             460
251  Met Gly Gly Lys Gln Gln Arg Arg Arg Asn Pro Gln Glu Gly Val Tyr
252  465             470             475             480
253  Asn Ala Leu Gln Lys Asp Lys Met Pro Glu Ala Tyr Ser Glu Ile Gly
254              485             490             495
255  Thr Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln
256              500             505             510
257  Asp Ser His Phe Gln Ala Val Gln Phe Gly Asn Arg Arg Glu Arg Glu
258      515             520             525
259  Gly Ser Glu Leu Thr Arg Thr Leu Gly Leu Arg Ala Arg Pro Lys Gly
260      530             535             540
261  Glu Ser Thr Gln Gln Ser Ser Gln Ser Cys Ala Ser Val Phe Ser Ile
262  545             550             555             560
263  Pro Thr Leu Trp Ser Pro Trp Pro Pro Ser Ser Ser Ser Gln Leu
264              565             570             575
```

266 (2) INFORMATION FOR SEQ ID NO: 5:
268   (i) SEQUENCE CHARACTERISTICS:
269   (A) LENGTH: 462 amino acids
270   (B) TYPE: amino acid
271   (C) STRANDEDNESS: single
272   (D) TOPOLOGY: linear
274   (ii) MOLECULE TYPE: protein
276   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
278  Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu
279  1               5               10              15
280  Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys
281              20              25              30
282  Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser
```

VERIFICATION SUMMARY   DATE: 02/11/2002
PATENT APPLICATION: US/09/939,537   TIME: 09:33:07

Input Set : N:\Crf3\RULE60\09939537.txt
Output Set: N:\CRF3\02112002\I939537.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:12 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:36 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]